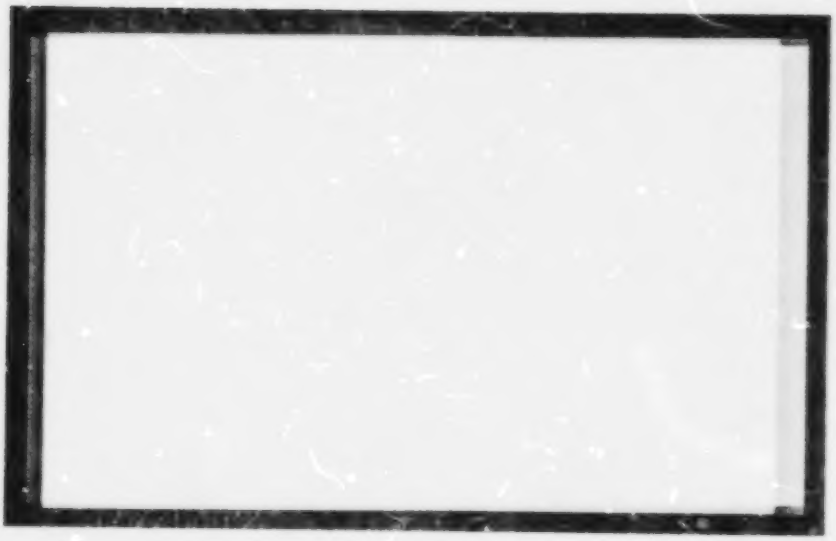
UNIVERSITÉ DE MONTRÉAL



**CENTRE DE RECHERCHE SUR LES TRANSPORTS**

CASE POSTALE 6128, SUCCURSALE A, MONTRÉAL, QUÉBEC, H3C 3J7

# LOCAL CONVERGENCE OF GAUSS-SEIDEL TECHNIQUES FOR NONLINEAR SPATIAL PRICE EQUILIBRIUM PROBLEMS

by

Jean-Pierre Crouzeix*
and Jacques A. Ferland**

* Département des mathématiques appliquées, Université de Clermont, France.

** Centre de recherche sur les transports et département d'informatique et de recherche opérationnelle, Université de Montréal, Québec, Canada.

This research was completed while Jacques A. Ferland was on sabbatical leave at the Université de Clermont and was supported by NSERC (IC-0218).

Cette étude a été publiée grâce à une subvention du fonds F.C.A.R. pour l'aide et le soutien à la recherche.

Université de Montréal

Centre de recherche sur les transports - Publication #496
Département d'informatique et
de recherche opérationnelle - Publication #587

Mars 1987

Dépôt légal - 1er trimestre 1987

## ABSTRACT

In this paper we study the local convergence of Gauss-Seidel type techniques introduced to handle the nonlinear spatial equilibrium problem. Although the Newton approach is embedded within these algorithms, their rate of convergence remains linear. Convergence ratios for these procedures are also analysed.



## RÉSUMÉ

Dans cette publication, nous étudions la convergence locale des méthodes de type Gauss-Seidel présentées pour traiter le problème d'équilibre spatial. Même si l'approche de Newton y est intégrée, il n'en reste pas moins que le taux de convergence n'est que linéaire. Les rapports de convergence de ces méthodes sont aussi contrastées.

## 1. INTRODUCTION

Gauss-Seidel type techniques were introduced in [1] to handle the nonlinear spatial equilibrium problem. Their global convergence was also analysed. The purpose of this paper is to study their local convergence and to compare their rates of convergence. The results are consistent with the numerical behavior of the techniques reported in [1].

Recall that the spatial price equilibrium problem is characterized by a set of supply nodes i, $1 \le i \le m$, and a set of demand nodes j, $1 \le j \le \ell$, for a single commodity. For each supply node i ($1 \le i \le m$), $d_i$ denotes the quantity supplied, and $S(d_i)$ the supply function specifying the price of the commodity at node i. For each demand node j ($1 \le j \le \ell$), $b_j$ denotes the quantity required, and $I_j(b_j)$ the inverse demand function specifying the cost of the commodity at node j. Furthermore, the commodity is transported from supply nodes to demand nodes; $t_{ij}$ denotes the quantity transported from i to j and $c_{ij}(t_{ij})$ is the transportation cost from i to j.

The price equilibrium conditions are as follows

$$S_i(d_i) + c_{ij}(t_{ij}) - I_j(b_j) \ge 0 \qquad 1 \le i \le m,\ 1 \le j \le \ell \qquad (1.1)$$

$$t_{ij}\,(S_i(d_i) + c_{ij}(t_{ij}) - I_j(b_j)) = 0 \qquad 1 \le i \le m,\ 1 \le j \le \ell \qquad (1.2)$$

$$d_i = \sum_{j=1}^{\ell} t_{ij} \qquad 1 \le i \le m \qquad (1.3)$$

$$b_j = \sum_{i=1}^{m} t_{ij} \qquad 1 \le j \le \ell \qquad (1.4)$$

$$t_{ij} \ge 0 \qquad 1 \le i \le m,\ 1 \le j \le \ell\ . \qquad (1.5)$$

Conditions (1.1) and (1.2) indicate that the flow of commodity $t_{ij}$ can be positive only if the price at the supply node added to the transportation cost is equal to the price at the demand node. Otherwise the flow is maintained at zero.

Now, assume the following hypothesis :

i) $S_i(d_i)$ is continuous, bounded, integrable, and monotone increasing, $1 \leq i \leq m$;

ii) $I_j(b_j)$ is continuous, bounded, integrable and monotone decreasing, $1 \leq j \leq \ell$;

iii) $c_{ij}(t_{ij})$ is continuous, bounded, integrable and monotone increasing, $1 \leq i \leq m$, $1 \leq j \leq \ell$.

It is easy to verify that the system (1.1) to (1.5) represents the Kuhn-Tucker conditions (necessary and sufficient) for the problem

$$\text{Min} \sum_{i=1}^{m} \int_0^{d_i} S_i(x)dx + \sum_{i=1}^{m} \sum_{j=1}^{\ell} \int_0^{t_{ij}} c_{ij}(x)dx - \sum_{j=1}^{\ell} \int_0^{b_j} I_j(y)dy$$

$$\text{subject to} \quad \sum_{j=1}^{\ell} t_{ij} = d_i \qquad 1 \leq i \leq m$$

$$\sum_{i=1}^{m} t_{ij} = b_j \qquad 1 \leq j \leq \ell$$

$$t_{ij} \geq 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq \ell \ .$$

Equivalently, this problem can be specified in terms of the variables $t_{ij}$ uniquely as follows :

$$\text{Min } h(t) = \sum_{i=1}^{m} \int_0^{\sum_{j=1}^{\ell} t_{ij}} S_i(x)dx + \sum_{i=1}^{m} \sum_{j=1}^{\ell} \int_0^{t_{ij}} c_{ij}(x)dx - \sum_{j=1}^{\ell} \int_0^{\sum_{i=1}^{m} t_{ij}} I_j(y)dy \qquad (1.6)$$

$$\text{subject to} \quad t_{ij} \geq 0 \qquad 1 \leq i \leq m,\ 1 \leq j \leq \ell.$$

This problem has a simple feasible domain reducing to the nonnegative orthant.

Convergence is analyzed in terms of the procedures applied to the more general problem

$$\underset{x \geq 0}{\text{Min}} \ f(x) \ . \qquad (1.7)$$

To establish the relation with problem (1.6), the assumptions required to complete the convergence analysis of the procedures are also interpreted in terms of this specific problem.

In Section 2, the 0-D pair decomposition technique is studied. This corresponds to a Gauss-Seidel iteractive procedure where the variables of (1.7) are modified one at a time. Then block decomposition procedures in which variables are modified by subsets are analysed in Section 3. Finally, in Section 4, the convergence study is shown to reduce to the analysis of the convergence of a sequence generated by the classical Gauss-Seidel method for the resolution of a linear system.

## 2. 0-D PAIR DECOMPOSITION

Referring to [1], the algorithm was defined in terms of the general problem

$$\min_{x>0} f(x). \tag{2.1}$$

The algorithm is an iteractive procedure generating a sequence $\{x^k\}$ of feasible solutions as follows : $x^{k+1} = x^k + z^k$ where $z_i^k (1<i<n)$ is an optimal solution to the problem

$$\min_{x_i^k+z_i>0} z_i \nabla_i f(x^k + \sum_{j=1}^{i-1} z_j^k e_j) + \frac{1}{2} z_i^2 \nabla_{ii} f(x^k + \sum_{j=1}^{i-1} z_j^k e_j) \tag{2.2}$$

where $\nabla_i f(x) = \frac{\partial f(x)}{\partial x_i}$ , $\nabla_{ij} f(x) = \frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ , and $e_i$ is the unit vector having its $i^{th}$ component equal to 1 and all the others equal to 0. Hence to determine $z^k$ we solve the following system obtained from the Kuhn-Tucker conditions for problems (2.2) :

$$
\begin{aligned}
&\nabla_1 f(x^k) + z_1^k \nabla_{11} f(x^k) - \mu_1^k = 0 \\
&\nabla_2 f(x^k + z_1^k e_1) + z_2^k \nabla_{22} f(x^k + z_1^k e_1) - \mu_2^k = 0 \\
&\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
&\nabla_n f(x^k + \sum_{j=1}^{n-1} z_j^k e_j) + z_n^k \nabla_{nn} f(x^k + \sum_{j=1}^{n-1} z_j^k e_j) - \mu_n^k = 0 \\
&\mu_1^k (x_1^k + z_1^k) = 0 \\
&\mu_2^k (x_2^k + z_2^k) = 0 \\
&\qquad\vdots \qquad\qquad\qquad\qquad\qquad\qquad\qquad \vdots \\
&\mu_n^k (x_n^k + z_n^k) = 0
\end{aligned}
\tag{2.3}
$$

where $\mu_i^k > 0$ and $x_i^k + z_i^k > 0$ for all $1 \leq i \leq n$.

To derive results on local convergence, we proceed similarly to Fiacco in [2] using the implicit function theorem applied to system (2.3) to derive information about $z^k = z(x^k)$ for $x^k$ in a neighborhood of $x^*$, an optimal solution of (2.1). The following assumptions concerning problem (2.1) are required :

$H_1$ : $x^*$ is the unique optimal solution of (2.1).

$H_2$ : $f$ is convex and three times continuously differentiable in a neighborhood of $x^*$.

$H_3$ : (strict complementarity) $\mu_j^* > 0$ when $x_j^* = 0$.

In view of the hypotheses required for specifying the spatial price equilibrium problem,

i) $H_2$ amounts to requiring that $S_i(d_i)$, $I_j(b_j)$ and $c_{ij}(t_{ij})$ be twice continuously differentiable for $1 \leq i \leq m$, $1 \leq j \leq \ell$.

ii) $H_3$ corresponds to the conditions $S_i(d_i^*) + c_{ij}(t_{ij}^*) - I_j(b_j^*) > 0$ when $t_{ij}^* = 0$, $1 \leq i \leq m$, $1 \leq j \leq \ell$. Indeed, it is easy to see that the optimal multiplier associated with the constraint $t_{ij} \geq 0$ is $S_i(d_i^*) + c_{ij}(t_{ij}^*) - I_j(b_j^*)$ for all $1 \leq i \leq m$, $1 \leq j \leq \ell$.

Let $H(z^k,\mu^k,x^k) = 0$ denote the system (2.3). Assumption $H_2$ implies that function H is jointly continuously differentiable in z, $\mu$, and x in a neighborhood of $(0,\mu^*,x^*)$. In particular, the 2nx2n Jacobian matrix $H'_{z,\mu}(0,\mu^*,x^*)$ is given by

$$H'_{z,\mu}(0,\mu^*,x^*) = \begin{bmatrix} \Delta & -I \\ U^* & X^* \end{bmatrix} \tag{2.4}$$

where I is the nxn identity matrix, $U^*$ and $X^*$ are nxn diagonal matrices, and $\Delta$ is an nxn lower triangular matrix such that

$$U^* = \begin{bmatrix} \mu_1^* & 0 & \dots & 0 \\ 0 & \mu_2^* & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mu_n^* \end{bmatrix}, \quad X^* = \begin{bmatrix} x_1^* & 0 & \dots & 0 \\ 0 & x_2^* & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & x_n^* \end{bmatrix}$$

$$\Delta = \begin{bmatrix} \nabla_{11} f(x^*) & 0 & \dots & 0 \\ \nabla_{21} f(x^*) & \nabla_{22} f(x^*) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \nabla_{n1} f(x^*) & \nabla_{n2} f(x^*) & \dots & \nabla_{nn} f(x^*) \end{bmatrix}. \tag{2.5}$$

Let $\Lambda = \{i : x_i^* > 0\} = \{i_1, i_2, \dots, i_q\}$ where $i_1 < i_2 < \dots < i_q$. We shall show that a necessary and sufficient condition for the matrix $H'_{z,\mu}(0,\mu^*,x^*)$ to be nonsingular is that $\nabla_{ii} f(x^*) \neq 0$ for all $i \in \Lambda$. (Recall that $\nabla_{ii} f(x^*) > 0$ for all i since f is convex.) Indeed, let a and $b \in R^n$ be such that

$$\begin{aligned} \Delta a - b &= 0 \\ U^* a - X^* b &= 0 \end{aligned} \tag{2.6}$$

In virtue of $H_3$, if $i \in \Lambda$, then $x_i^* > 0$, $\mu_i^* = 0$, and consequently $b_i = 0$; whereas, if $i \notin \Lambda$, then $x_i^* = 0$, $\mu_i^* > 0$, and consequently $a_i = 0$. Define the vector $v \in R^n$ and the nxn lower triangular matrix M as follows :

$$v_i = \begin{cases} a_i & \text{if } i \in \Lambda \\ b_i & \text{if } i \notin \Lambda \end{cases}$$

$$m_{ij} = \begin{cases} \Delta_{ij} & \text{if } j<i \text{ and } j \in \Lambda \\ -1 & \text{if } i=j \text{ and } j \notin \Lambda \\ 0 & \text{otherwise.} \end{cases}$$

Then system (2.6) reduces to $Mv = 0$, and $H'_{z,\mu}(0,\mu^*,x^*)$ is nonsingular if and only if M is nonsingular. Hence the result follows.

For the global convergence analysis in [1], the stronger requirement that $\nabla_{ii}f(x) > 0$ for all $x>0$, $1<i<n$, was needed. Note that in terms of the spatial price equilibrium problem, this corresponds to requiring that $S'(d_i) + c'_{ij}(t_{ij}) - I'_j(b_j) > 0$, $1<i<m$, $i<j<\ell$.

If $\nabla_{ii}f(x^*) > 0$ for all $1<i<n$, then the assumptions of the implicit function theorem are satisfied for function H at the point $(0,\mu^*,x^*)$. Hence there exists a neighborhood V of $x^*$ and a unique continuously differentiable function $[z(x),\mu(x)]$ on V with $z(x^*) = 0$, $\mu(x^*) = \mu^*$ and $H(z(x),\mu(x),x) = 0$ for all $x \in V$. Furthermore, since $\mu_i^* = \mu_i(x^*) > 0$ for all $i \notin \Lambda$, $x_i^* + z_i(x^*) > 0$ for all $i \in \Lambda$, and the functions $\mu$ and z are continuous, it follows that there exists a neighborhood W of $x^*$, $W \subset V$, such that for all $x \in W$,

$$\mu_i(x) > 0 \qquad \text{for all } i \notin \Lambda$$

$$x_i + z_i(x) > 0 \qquad \text{for all } i \in \Lambda.$$

Finally, since, for all $x \in W$, $\mu_i(x)(x_i + z_i(x)) = 0$ for all i, one has

$$\mu_i(x) = 0 \qquad \text{for all } i \in \Lambda$$

$$x_i + z_i(x) = 0 \qquad \text{for all } i \notin \Lambda.$$

Referring to the global convergence analysis, it is known that $x^{k-1} \in W$ for k large enough. Since $\mu_i^{k-1} = \mu_i(x^{k-1})$, $z_i^{k-1} = z_i(x^{k-1})$ and $x^k = x^{k-1} + z^{k-1}$, then

$$z_i^k = x_i^k = 0 \qquad \text{for all } i \notin \Lambda$$

$$\mu_i^k = 0 \qquad \text{for all } i \in \Lambda.$$

Hence deleting the rows not in $\Lambda$ from system (2.3), it follows that

$$\begin{aligned}
&\nabla_{i_1} f(x^k) + z_{i_1}^k \nabla_{i_1 i_1} f(x^k) = 0 \\
&\nabla_{i_2} f(x^k + z_{i_1}^k e_{i_1}) + z_{i_2}^k \nabla_{i_2 i_2} f(x^k + z_{i_1}^k e_{i_1}) = 0 \qquad (2.7) \\
&\vdots \qquad\qquad\qquad \vdots \qquad\qquad \vdots \\
&\nabla_{i_q} f(x^k + \sum_{j=1}^{q-1} z_{i_j}^k e_{i_j}) + z_{i_q}^k \nabla_{i_q i_q} f(x^k + \sum_{j=1}^{q-1} z_{i_j}^k e_{i_j}) = 0.
\end{aligned}$$

Let $\tilde{x} = [x_{i_1}, x_{i_2}, \ldots, x_{i_q}]^T$ and $\tilde{z} = [z_{i_1}, z_{i_2}, \ldots, z_{i_q}]^T$. Then (2.7) is summarized as

$$\tilde{H}(\tilde{z}^k, \tilde{x}^k) = 0 \ .$$

The implicit function theorem can also be applied to $\tilde{H}$ at the point $(0, \tilde{x}^*)$. Hence, there exists a continuously differentiable function $\tilde{z}(\tilde{x})$ with $\tilde{z}(\tilde{x}^*)=0$ and $\tilde{H}(\tilde{z}(\tilde{x}), \tilde{x}) = 0$ for all $\tilde{x}$ in a neighborhood of $\tilde{x}^*$. Furthermore, since

$$\tilde{H}'_{\tilde{z}}(0, \tilde{x}^*)\tilde{z}'(\tilde{x}^*) + \tilde{H}'_{\tilde{x}}(0, \tilde{x}^*) = 0 \ ,$$

it follows that

$$\tilde{z}'(\tilde{x}^*) = -\tilde{\Delta}^{-1}\tilde{R} \qquad (2.8)$$

where $\tilde{\Delta}$ is obtained from $\Delta$, and $\tilde{R}$ from the Hessian matrix of f by deleting the rows and columns not in $\Lambda$.

Now, we can complete the study of the local convergence of the procedure. Assume that $x^k \in W$. Then

$$\tilde{x}^{k+1} - \tilde{x}^* = \tilde{x}^k - \tilde{x}^* + \tilde{z}(\tilde{x}^k)$$

$$\tilde{x}^{k+1} - \tilde{x}^* = \tilde{x}^k - \tilde{x}^* + \tilde{z}'(\tilde{x}^*)(\tilde{x}^k - \tilde{x}^*) + \varepsilon_k \|\tilde{x}^k - \tilde{x}^*\|$$

where $\varepsilon_k$ tends to zero when $\tilde{x}^k \to x^*$. It follows from (2.7) that

$$\tilde{x}^{k+1} - \tilde{x}^* = (\tilde{I} - \tilde{\Delta}^{-1}\tilde{R})(\tilde{x}^k - \tilde{x}^*) + \varepsilon_k \|\tilde{x}^k - \tilde{x}^*\| \quad .$$

Since $\|\tilde{x}^k - \tilde{x}^*\| = \|x^k - x^*\|$ for the euclidean norm or $\|\cdot\|_1$, it follows then that the procedure converges linearly with the convergence ratio $\|\tilde{I} - \tilde{\Delta}^{-1}\tilde{R}\|$ .

Furthermore, to establish a relation with the classical Gauss-Seidel method, define the diagonal matrix $\tilde{D}$ and the strictly lower triangular matrix $\tilde{L}$ as follows :

$$\tilde{D} = \begin{bmatrix} \nabla_{i_1 i_1} f(x^*) & 0 & \dots & 0 \\ 0 & \nabla_{i_2 i_2} f(x^*) & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & & \nabla_{i_q i_q} f(x^*) \end{bmatrix} \qquad (2.9)$$

$$\tilde{L} = -\begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ \nabla_{i_2 i_1} f(x^*) & 0 & 0 & \dots & 0 \\ \nabla_{i_3 i_1} f(x^*) & \nabla_{i_3 i_2} f(x^*) & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \nabla_{i_q i_1} f(x^*) & \nabla_{i_q i_2} f(x^*) & \nabla_{i_q i_3} f(x^*) & \dots & \nabla_{i_q i_{q-1}} f(x^*) \quad 0 \end{bmatrix} . \qquad (2.10)$$

Then

$$\tilde{R} = \tilde{D} - \tilde{L} - \tilde{L}^T$$

$$\tilde{\Delta} = \tilde{D} - \tilde{L} \quad .$$

Hence, it follows that for k large enough so that $x^k \in W$

$$(\tilde{x}^{k+1} - \tilde{x}^*) = (\tilde{D} - \tilde{L})^{-1} \tilde{L}^T (\tilde{x}^k - \tilde{x}^*) . \qquad (2.11)$$

We analyse the convergence of this process in section 4.

## 3. BLOCK DECOMPOSITION

The block decomposition algorithms introduced in [1] are active sets methods in the sense that the active constraints are identified in a neighborhood of the optimal solution. It follows that to analyse the local convergence of these methods is equivalent to analysing the convergence of the underlying procedure which solves the unconstrained problem

$$\text{Min } g(\tilde{x}) \qquad (3.1)$$

where $g(\tilde{x}) = f(x)$ with $x_i = 0$ if $i \notin \Lambda$ and $\tilde{x} = [x_{i_1}, x_{i_2}, \ldots, x_{i_q}]^T$ .

Suppose we have the vector $\tilde{x} = [\tilde{x}_1^T, \tilde{x}_2^T, \ldots, \tilde{x}_\tau^T]^T \in R^q$ where $\tilde{x}_s \in R^{n_s}$ with $\sum_{s=1}^{\tau} n_s = q$. Partition analogously the gradient and the Hessian matrix:

$$\nabla g(\tilde{x}) = [\tilde{\nabla}_1 g(\tilde{x})^T, \tilde{\nabla}_2 g(\tilde{x})^T, \ldots, \tilde{\nabla}_\tau g(\tilde{x})^T]^T$$

$$\nabla^2 g(\tilde{x}) = \begin{bmatrix} \tilde{\nabla}_{11} g(\tilde{x}) & \tilde{\nabla}_{12} g(\tilde{x}) & \ldots & \tilde{\nabla}_{1\tau} g(\tilde{x}) \\ \tilde{\nabla}_{21} g(\tilde{x}) & \tilde{\nabla}_{22} g(\tilde{x}) & \ldots & \tilde{\nabla}_{2\tau} g(\tilde{x}) \\ \vdots & \vdots & & \vdots \\ \tilde{\nabla}_{\tau 1} g(\tilde{x}) & \tilde{\nabla}_{\tau 2} g(\tilde{x}) & \ldots & \tilde{\nabla}_{\tau\tau} g(\tilde{x}) \end{bmatrix} .$$

For k large enough, $\tilde{x}^k$ 0 is in a neighborhood of $\tilde{x}^*$ 0, and the underlying procedure generates a sequence $\{\tilde{x}^k\}$ such that $\tilde{x}^{k+1} = \tilde{x}^k + \tilde{z}^k$ where for $1 \le s \le \tau$, $\tilde{z}_s^k$ is an optimal solution of the problem

$$\text{Min } \tilde{\nabla}_s g(\tilde{x}^k + \sum_{j=1}^{s-1} \tilde{e}_j \tilde{z}_j^k)^T \tilde{z}_s + \frac{1}{2} \tilde{z}_s^T \tilde{\nabla}_{ss} g(\tilde{x}^k + \sum_{j=1}^{s-1} \tilde{e}_j \tilde{z}_j^k) \tilde{z}_s \qquad (3.2)$$

where $[\tilde{e}_1, \tilde{e}_2, \ldots, \tilde{e}_\tau] = I_q$.

Hence to determine $\tilde{z}^k$ we solve the following system obtained from the optimality conditions for problems (3.2):

$$\begin{aligned}
&\tilde{\nabla}_1 g(\tilde{x}^k) + \tilde{\nabla}_{11}(g(\tilde{x}^k)\, \tilde{z}_1^k = 0 \\
&\tilde{\nabla}_2 g(\tilde{x}^k + \tilde{e}_1 \tilde{z}_1^k) + \tilde{\nabla}_{22} g(\tilde{x}^k + \tilde{e}_1 \tilde{z}_1^k)\, \tilde{z}_2^k = 0 \qquad (3.3) \\
&\quad \vdots \qquad\qquad\qquad \vdots \qquad\qquad \vdots \\
&\tilde{\nabla}_\tau g(\tilde{x}^k + \sum_{j=1}^{\tau-1} \tilde{e}_j \tilde{z}_j{}^k) + \tilde{\nabla}_{\tau\tau} g(\tilde{x}^k + \sum_{j=1}^{\tau-1} \tilde{e}_j \tilde{z}_j^k)\, \tilde{z}_\tau^k = 0
\end{aligned}$$

As in the preceding section, we derive results on local convergence using the implicit function theorem. Let $\tilde{H}(\tilde{z}^k, \tilde{x}^k) = 0$ denote the system (3.3). Assume that $H_1$ and $H_2$ hold. Furthermore, under the hypothesis that $\tilde{\nabla}_{ss} g(x^*)$ is nonsingular $1 \le s \le \tau$, it follows that

$$\tilde{\Delta}_\tau = \begin{bmatrix}
\tilde{\nabla}_{11} g(\tilde{x}^*) & 0 & \ldots & 0 \\
\tilde{\nabla}_{21} g(\tilde{x}^*) & \tilde{\nabla}_{22} g(\tilde{x}^*) & \ldots & 0 \\
\vdots & \vdots & & \vdots \\
\tilde{\nabla}_{\tau 1} g(\tilde{x}^*) & \tilde{\nabla}_{\tau 2} g(\tilde{x}^*) & \ldots & \tilde{\nabla}_{\tau\tau} g(\tilde{x}^*)
\end{bmatrix} \qquad (3.4)$$

is invertible. Hence the implicit function theorem implies that in the neighborhood of $\tilde{x}^*$ there exists a unique continuously differentiable function $\tilde{z}(\tilde{x})$ satisfying (3.3) with $\tilde{z}(\tilde{x}^*) = 0$.

To derive $\tilde{z}'(\tilde{x}^*)$, we solve the system

$$\tilde{H}'_{\tilde{z}}(0, \tilde{x}^*)\, \tilde{z}'(\tilde{x}^*) + \tilde{H}'_{\tilde{x}}(0, \tilde{x}^*) = 0 \quad .$$

This system is written explicitly as

$$\tilde{\nabla}_\tau \tilde{z}'(\tilde{x}^*) = -\nabla^2 g(\tilde{x}^*)$$

and (3.5)

$$\tilde{z}'(\tilde{x}^*) = \tilde{\Delta}_\tau^{-1} \nabla^2 g(x^*) \quad .$$

Recall that $\nabla^2 g(\tilde{x})$ is obtained from the Hessian of f by deleting the rows and columns not in $\Lambda$ so that $\nabla\, g(\tilde{x}^*) = \tilde{R}$. Hence, using an argument similar to Section 2, it follows that the procedure converges linearly with the convergence ratio $\|\tilde{I} - \tilde{\Delta}_\tau^{-1}\tilde{R}\|$.

Furthermore, define the block diagonal matrix $\tilde{D}_\tau$ and the strictly lower triangular matrix $\tilde{L}_\tau$ as follows

$$\tilde{D}_\tau = \begin{bmatrix} \tilde{\nabla}_{11} g(\tilde{x}^*) & 0 & \dots & 0 \\ 0 & \tilde{\nabla}_{22} g(\tilde{x}^*) & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{\nabla}_{\tau\tau} g(x^*) \end{bmatrix} \tag{3.6}$$

$$\tilde{L}_\tau = - \begin{bmatrix} 0 & 0 & & 0 & . \; . \; . & 0 \\ \tilde{\nabla}_{21} g(\tilde{x}^*) & 0 & \dots & 0 & . \; . \; . & 0 \\ \tilde{\nabla}_{31} g(\tilde{x}^*) & \tilde{\nabla}_{32} g(\tilde{x}^*) & \dots & 0 & . \; . \; . & 0 \\ \vdots & \vdots & & \vdots & . \; . \; . & \vdots \\ \tilde{\nabla}_{\tau 1} g(\tilde{x}^*) & \tilde{\nabla}_{\tau 2} g(\tilde{x}^*) & \dots & \tilde{\nabla}_{\tau 3} g(\tilde{x}^*) \; \dots & \tilde{\nabla}_{\tau\tau-1} g(\tilde{x}^*) & 0 \end{bmatrix} . \tag{3.7}$$

Then

$$\tilde{R} = \tilde{D}_\tau - \tilde{L}_\tau - \tilde{L}_\tau^T$$

$$\tilde{\Delta}_\tau = \tilde{D}_\tau - \tilde{L}_\tau \quad .$$

Hence it follows that in the neighborhood of the solution

$$(\tilde{x}^{k+1} - \tilde{x}^*) = (\tilde{D}_\tau - \tilde{L}_\tau)^{-1} \tilde{L}_\tau^T (\tilde{x}^k - x^*) \ . \qquad (3.8)$$

## 4. CONVERGENCE RATIOS

First note that in the neighborhood of the optimal solution $\|x^{k+1} - x^*\| = \|\tilde{x}^{k+1} - \tilde{x}^*\|$ for the euclidean norm and $\|\cdot\|_1$. Hence, in both cases, the convergence study of the procedures reduces to analysing the convergence to zero of a sequence generated by an iteractive process of the type

$$\varepsilon^{k+1} = (D-L)^{-1} L^T \varepsilon^k \qquad (4.1)$$

where $\tilde{R} = D-L-L^T$ and D and L are specified by relations (2.9) and (2.10) or relations (3.6) and (3.7). Furthermore D and $\tilde{R}$ are symmetric, D is positive definite and (D-L) is nonsingular.

Relation (4.1) corresponds to the classical Gauss-Seidel method for the resolution of a linear system of the type $\tilde{R}x=b$. Then, it is well known tnat the method converges linearly if matrix $\tilde{R}$ is positive definite (see, [4, corollary 1, p. 78]). It is also known [4, exercise 1, p. 80-81] that the block decomposition does not necessarily generate an improvement over the corresponding point method. The improvement, if any, clearly depends on the partition. Nevertheless, the numerical results in [1] indicate that an appropriate partition can accelerate the procedure.

Finally it is interesting to note that although these Gauss-Seidel techniques embed the Newton approach, they do not take on the quadratic convergence of this latter approach.

## REFERENCES

[1] FERLAND, J.A., ZUBIETA, L. and SPIESS, H., "Nonlinear Spatial Price Equilibrium", Publication #477. Centre de recherche sur les transports, Université de Montréal, (1986).

[2] FIACCO, A.V., "Sensitivity Analysis for Nonlinear Programming Using Penalty Methods", Mathematical Programming, 10, 287-311 (1976).

[3] ORTHEGA, J.M. and RHEINBOLDT, W.C., "Iterative Solution of Nonlinear Equations in Several Variables, Prentice-Hall, Englewood Cliffs, New Jersey (1970).

[4] VARGA, R.S., "Matrix Iterative Analysis", Prentice-Hall, Englewood Cliffs, New Jersey (1962).